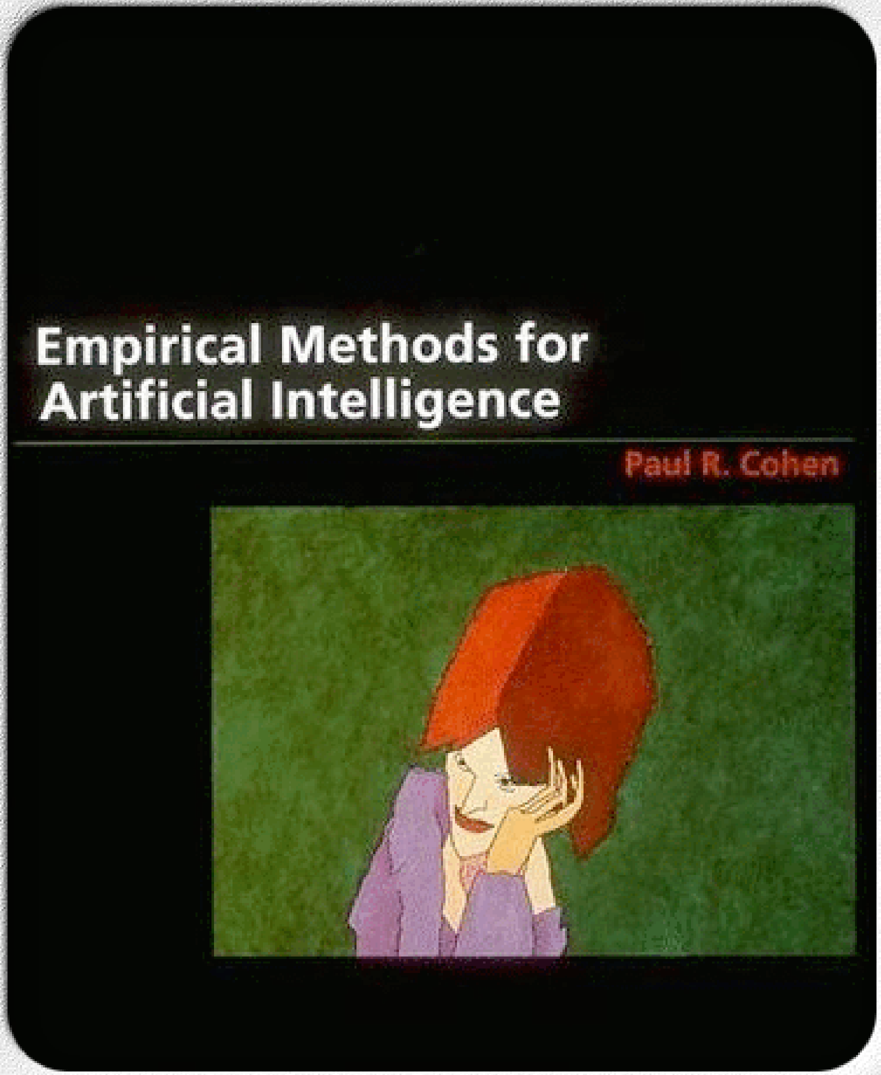
Optimization weekly meeting
May 4, 2006



Chapter 4:
Hypothesis Testing and Estimation

IRIDIA - ULB
Max MANFRIN

Outline

what we did in part 1 and what we have to do in part 2

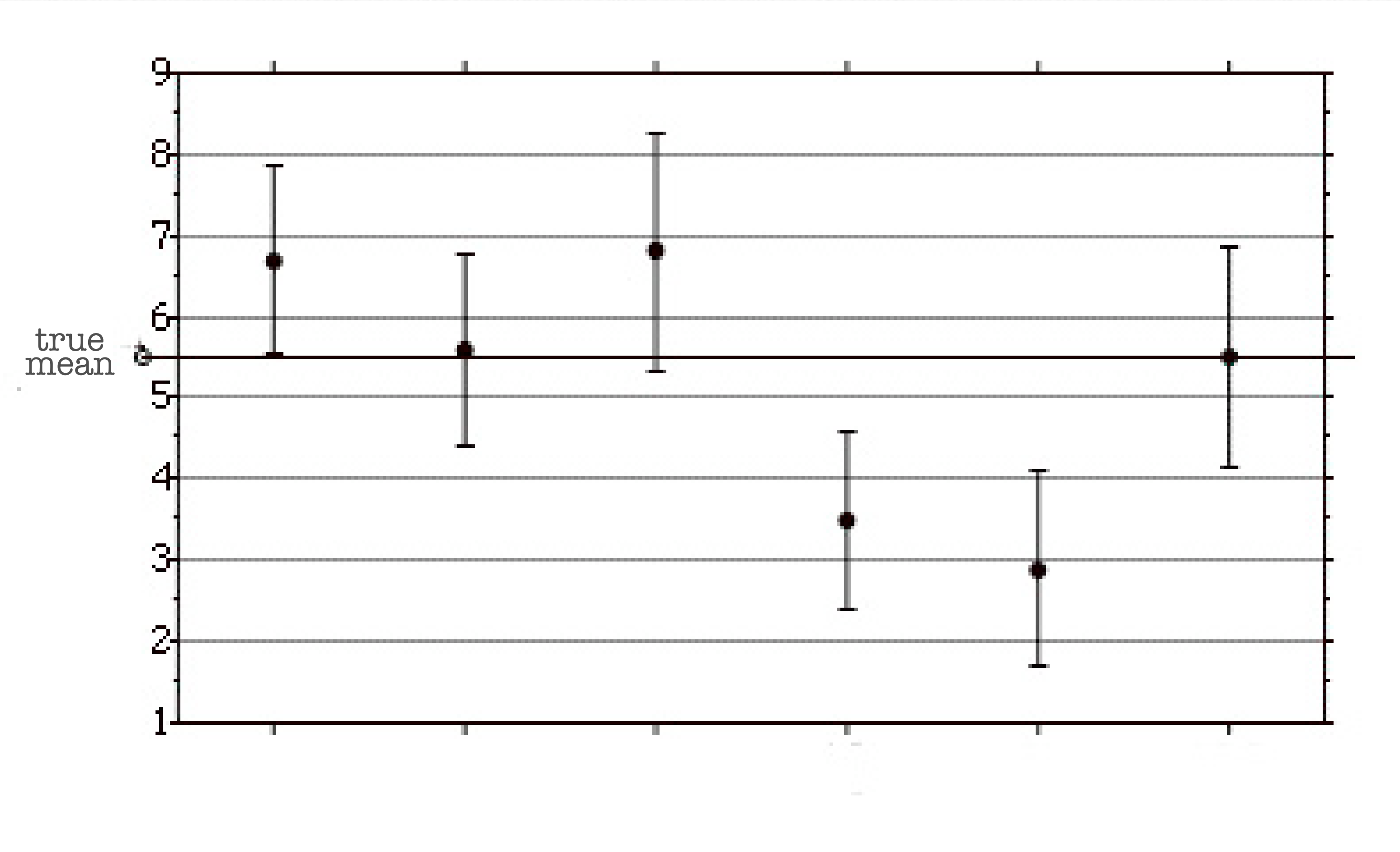
- Statistical Inference
- Introduction to Hypothesis Testing
- Sampling Distributions and Hypothesis Testing Strategy
- Test of Hypotheses about Means
- Hypothesis about Correlations
- Parameter Estimation and Confidence Intervals
- How big Should Samples Be?
- Errors
- Power Curves and How to Get Them

Outline

part 2

- Parameter Estimation and Confidence Intervals
- How big Should Samples Be?
- Errors
- Power Curves and How to Get Them

A **confidence interval** gives an estimated range of values (calculated from a given set of sample data) which is likely to include an unknown population parameter like μ_x, σ^2, σ

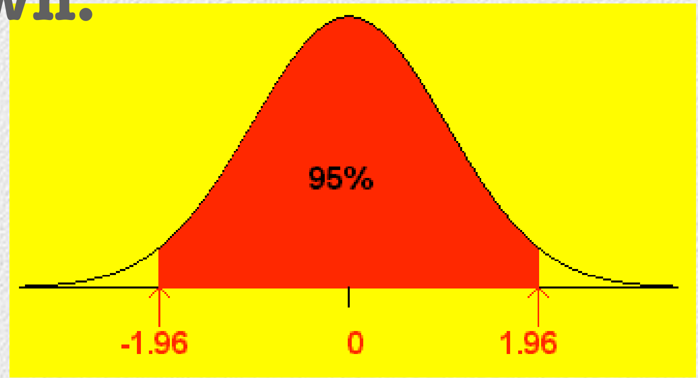


Java applet to demo confidence intervals for means

http://www.math.csusb.edu/faculty/stanton/m262/confidence_means/confidence_means.html

Sec 4.6

Confidence intervals for μ when σ is known:



$\bar{x} = \mu \pm 1.96\sigma_{\bar{x}}$ for 95% of the means $\bar{x}$

If $\epsilon = 1.96\sigma_{\bar{x}}$ the confidence interval $\bar{x} \pm \epsilon$ will contain μ in 95% of the samples we draw.

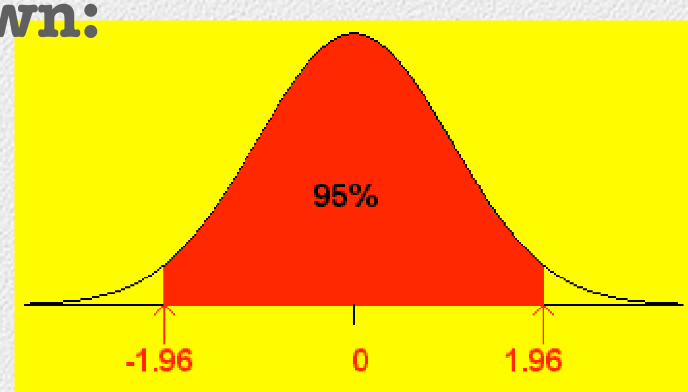
Confidence intervals for μ when σ is unknown:

$\bar{x} = \mu \pm t_{0.025}\hat{\sigma}_{\bar{x}}$ for 95% of the means $\bar{x}$

If $\epsilon = t_{0.025}\hat{\sigma}_{\bar{x}}$ the confidence interval $\bar{x} \pm \epsilon$ will contain μ in 95% of the samples we draw.

Sec 4.6

Confidence intervals for μ when σ is known:



$\bar{x} = \mu \pm 1.96\sigma_{\bar{x}}$ for 95% of the means $\bar{x}$

If $\epsilon = 1.96\sigma_{\bar{x}}$ the confidence interval $\bar{x} \pm \epsilon$ will contain μ in 95% of the samples we draw.

Confidence intervals for μ when σ is un

$\bar{x} = \mu \pm t_{0.025}\hat{\sigma}_{\bar{x}}$ for 95% of the means $\bar{x}$

If $\epsilon = t_{0.025}\hat{\sigma}_{\bar{x}}$ the confidence interval $\bar{x}$ contain μ in 95% of the samples we draw.

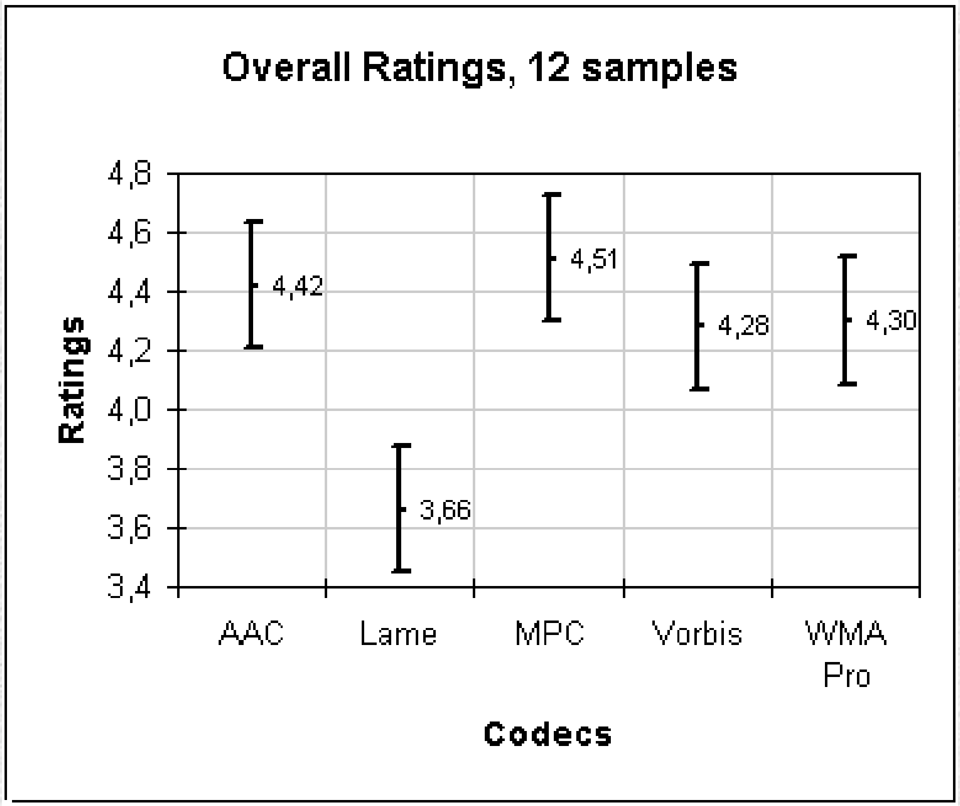
Confidence intervals

Because confidence intervals depend on standard errors, they are affected by sample size.

Sec 4.6

In general, one shouldn't use confidence intervals to test hypotheses if the confidence intervals overlap.

Confidence interval for means are estimated from standard errors of individual means, while two-sample t-test relies on standard error of the difference between the means.
Different sampling distribution and standard errors are involved.



Sec 4.6

Outline

part 2

- ~~Parameter Estimation and Confidence Intervals~~
- How big Should Samples Be?
- Errors
- Power Curves and How to Get Them

In parameter estimation if we increase the sample size and the sample variance doesn't increase (a reasonable assumption), then the confidence interval narrows.
(Our confidence in the estimate doesn't change, but the range of the estimate reduces)

In hypothesis testing, by contrast, the only thing that changes when increasing the sample size is our confidence in the conclusion.
(so if our sample is large enough to make us confident, nothing is gained by having a larger sample)

Sec 4.7

In parameter e... timation if we increase the sample size and the sample ... increase (a reasonable assumption), t...

(Our confidence i... the estimate reduces)

In hypothesis ... hat changes when increa... ence in the conclusion.

(so if our sample ... gained by having a larger sample)

Parameter estimation

Samples should be as large as we can afford

Hypothesis Testing

Samples should be no larger than required to show an effect

Sec 4.7

Samples can be too big for hypothesis testing.

Any real statistical effect can be boosted to significance by increasing N.

$$\bar{x} = \mu \pm t_{0.025}\hat{\sigma}_{\bar{x}}$$

(this is because the standard error of any statistic is reduced by increasing N)

One can obtain $\bar{x}_1 \neq \bar{x}_2$ in a statistical significant way by increasing N.

The whole point of statistical inference is to go beyond sample results to say something about populations, to predict future results.

$\boxed{\text{Sec 4.7}}$

Outline

part 2

- ~~Parameter Estimation and Confidence Intervals~~
- ~~How big Should Samples Be?~~
- Errors
- Power Curves and How to Get Them

In hypothesis testing:

	reject H_0	don't reject H_0
if H_0 true	type I error	✓
if H_1 true	✓	type II error

Sec 4.8

In hypothesis testing:

	reject H_0	don't reject H_0
if H_0 true		
if H_1 true		

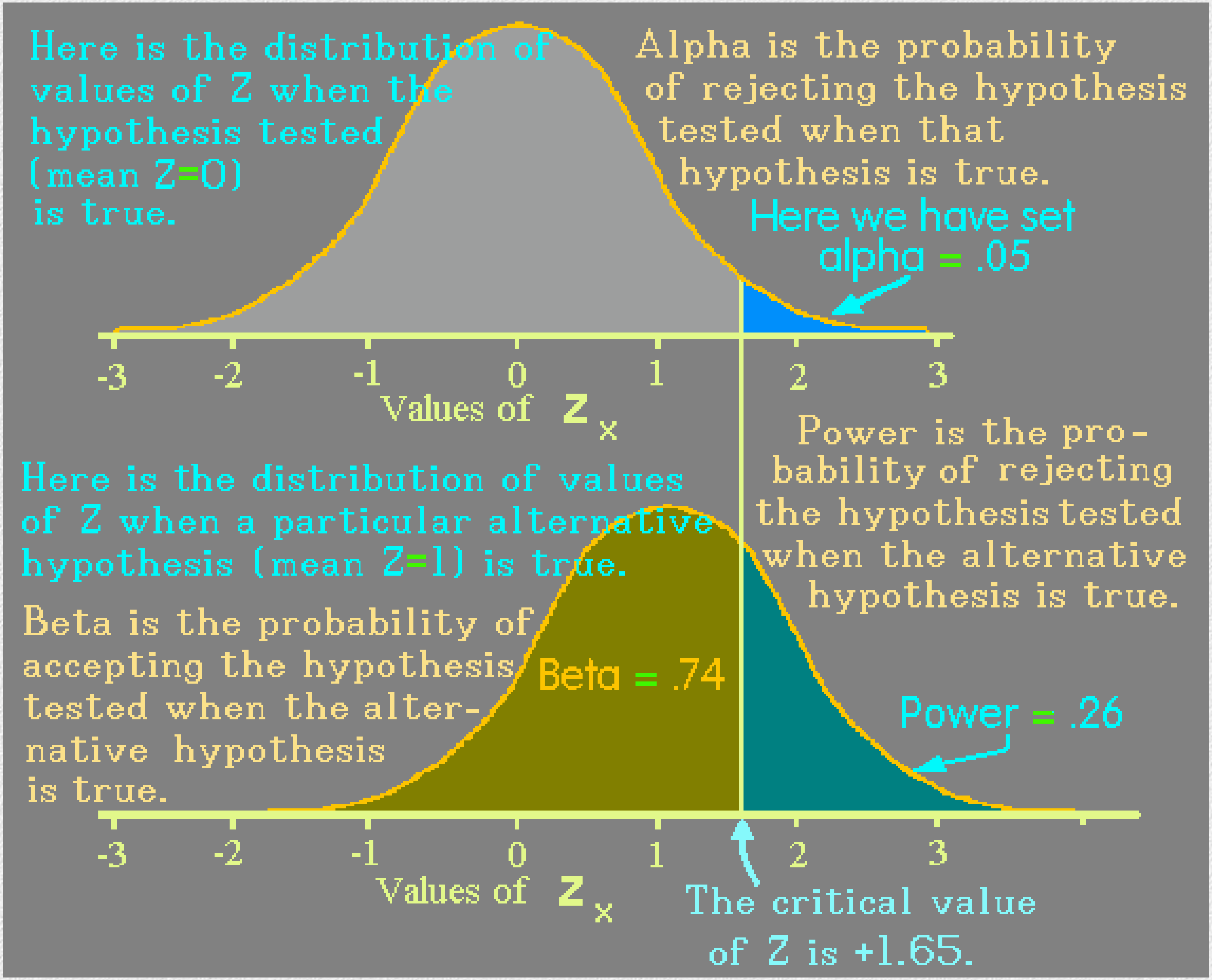
Type I error

$\alpha = \Pr(\text{ type I error}) = \Pr(\text{Reject } H_0 \mid H_0 \text{ is true})$

Type II error

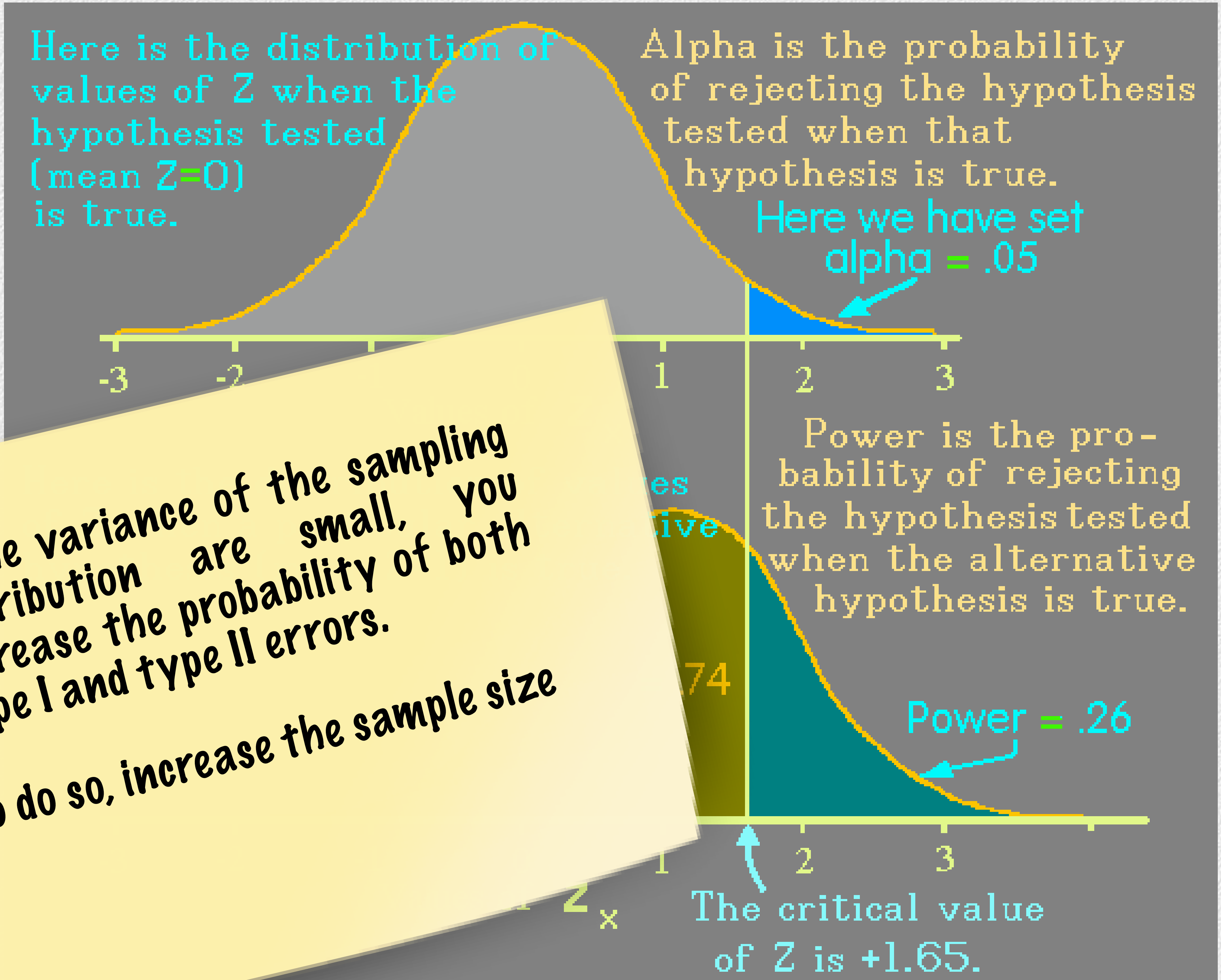
$\beta = \Pr(\text{ type II error}) = \Pr(\text{Fail to reject } H_0 \mid H_0 \text{ is false})$

Sec 4.8



http://www.animatedsoftware.com/statglos/sgbeta.htm

Sec 4.8



If the variance of the sampling distribution are small, you decrease the probability of both type I and type II errors.

To do so, increase the sample size

http://www.animatedsoftware.com/statglos/sgbeta.htm

Sec 4.8

Outline

part 2

- ~~Parameter Estimation and Confidence Intervals~~
- ~~How big Should Samples Be?~~
- ~~Errors~~
- Power Curves and How to Get Them

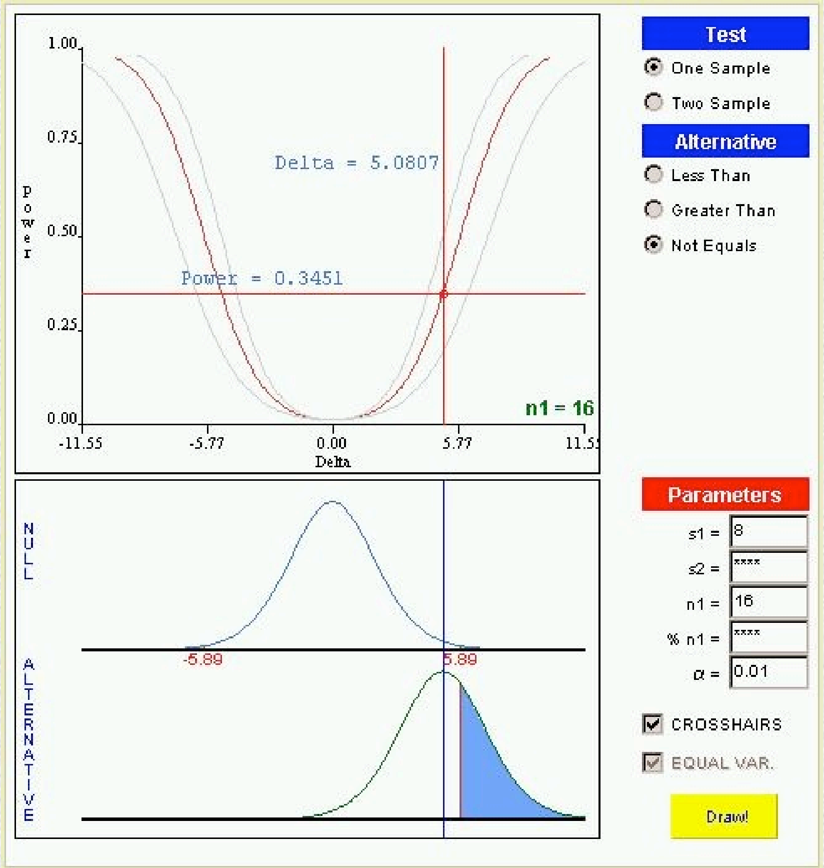
13

The power of a test depends on:

- the α level of the test;
- the degree of separation between H_0 and H_1 distributions;
- the variances of the populations from which samples are drawn;
- sample size N

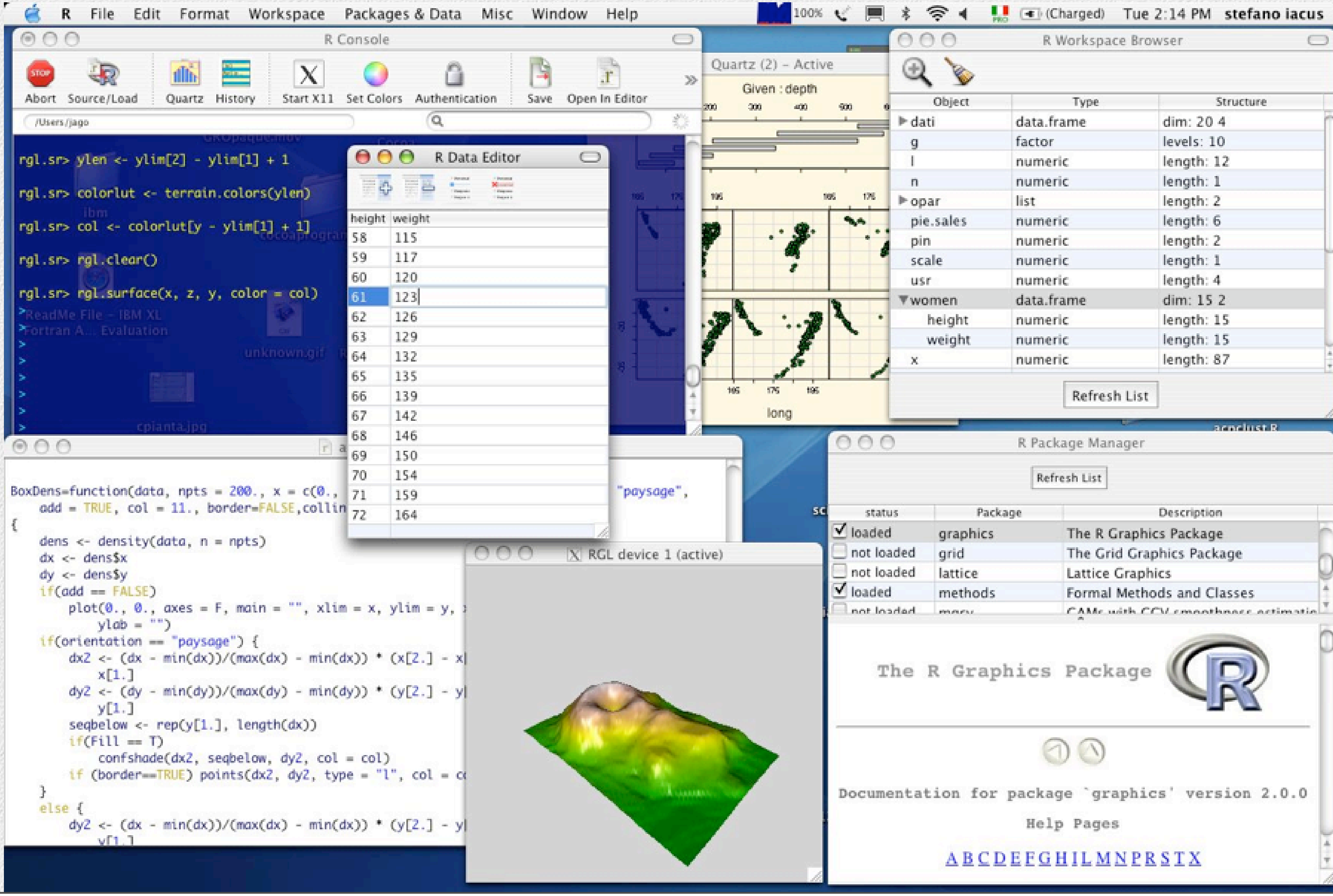
Power curves plot power against changes in one of these factors.

http://www.amstat.org/publications/jse/v11n3/anderson-cook.html



Sec 4.9

Demo



R File Edit Format Workspace Packages & Data Misc Window Help
100% (Charged) Tue 2:14 PM stefano iacus
R Console
Abort Source/Load Quartz History Start X11 Set Colors Authentication Save Open In Editor
/Users/jago
rgl.sr> ylen <- ylim[2] - ylim[1] + 1
rgl.sr> colorlut <- terrain.colors(ylen)
rgl.sr> col <- colorlut[y - ylim[1] + 1]
rgl.sr> rgl.clear()
rgl.sr> rgl.surface(x, z, y, color = col)
R Data Editor
height weight
58 115
59 117
60 120
61 123
62 126
63 129
64 132
65 135
66 139
67 142
68 146
69 150
70 154
71 159
72 164
Quartz (2) – Active
Given : depth
long
R Workspace Browser
Object Type Structure
dati data.frame dim: 20 4
g factor levels: 10
l numeric length: 12
n numeric length: 1
opar list length: 2
pie.sales numeric length: 6
pin numeric length: 2
scale numeric length: 1
usr numeric length: 4
women data.frame dim: 15 2
height numeric length: 15
weight numeric length: 15
x numeric length: 87
Refresh List
BoxDens=function(data, npts = 200., x = c(0., "paysage",
add = TRUE, col = 11., border=FALSE,collin
{
dens <- density(data, n = npts)
dx <- dens$x
dy <- dens$y
if(add == FALSE)
plot(0., 0., axes = F, main = "", xlim = x, ylim = y,
ylab = "")
if(orientation == "paysage") {
dx2 <- (dx - min(dx))/(max(dx) - min(dx)) * (x[2.] - x
x[1.]
dy2 <- (dy - min(dy))/(max(dy) - min(dy)) * (y[2.] - y
y[1.]
seqbelow <- rep(y[1.], length(dx))
if(Fill == T)
confshade(dx2, seqbelow, dy2, col = col)
if (border==TRUE) points(dx2, dy2, type = "l", col = c
}
else {
dy2 <- (dx - min(dx))/(max(dx) - min(dx)) * (y[2.] - y
RGL device 1 (active)
R Package Manager
Refresh List
status Package Description
loaded graphics The R Graphics Package
not loaded grid The Grid Graphics Package
not loaded lattice Lattice Graphics
loaded methods Formal Methods and Classes
The R Graphics Package
Documentation for package `graphics' version 2.0.0
Help Pages
A B C D E F G H I L M N P R S T X